WO 2005/103296

<130> VTOB.302VPC

PCT/US2005/010733

L

SEQUENCE LISTING

<110> Vector Tobacco, Ltd.

<120> GLOBAL GENE EXPRESSION ANALYSIS OF HUMAN BRONCHIAL EPITHELIAL CELLS EXPOSED TO CIGARETTE SMOKE, SMOKE CONDENSATES, OR COMPONENTS THEREOF

.

```
<140> Unknown
<141> 2005-03-29
<150> 60/557,929
<151> 2004-03-30
<160> 5
<170> FastSEQ for Windows Version 4.0
<210>1
<211> 5688
<212> DNA
<213> Artificial Sequence
<223> Full-length QTPase RNAi construct
<400> 1
ctcgaggatc taaattgtga gttcaatctc ttccctattg gattgattat cctttctttt 60
cttccaattt gtgtttcttt ttgcctaatt tattgtgtta tcccctttat cctattttgt 120
ttctttactt atttatttgc ttctatgtct ttgtacaaag atttaaactc tatggcacat 180
attttaaagt tgttagaaaa taaattcttt caagattgat gaaagaactt tttaattgta 240
gatatttcgt agattttatt ctcttactac caatataacg cttgaattga cgaaaatttg 300
tgtccaaata tctagcaaaa aggtatccaa tgaaaatata tcatatgtga tcttcaaatc 360
ttgtgtctta tgcaagattg atactttgtt caatggaaga gattgtgtgc atatttttaa 420
aatttttatt agtaataaag attctatata gctgttatag agggataatt ttacaaagaa 480
cactataaat atgattgttg ttgttagggt gtcaatggtt cggttcgact ggttatttta 540
taaaatttgt accataccat ttttttcgat attctatttt gtataaccaa aattagactt 600
ttcgaaatcg tcccaatcat gtcggtttca cttcggtatc ggtaccgttc ggttaatttt 660
cattttttt taaatgtcat taaaattcac tagtaaaaat agaatgcaat aacatacgtt 720
cttttatagg acttagcaaa agctctctag acatttttac tgtttaaagg ataatgaatt 780
aaaaaacatg aaagatggct agagtataga tacacaacta ttcgacagca acgtaaaaga 840
aaccaagtaa aagcaaagaa aatataaatc acacgagtgg aaagatatta accaagttgg 900
gattcaagaa taaagtctat attaaatatt caaaaagata aatttaaata atatgaaagg 960
aaacatatte aatacattgt agtttgetae teataatege tagaataett tgtgeettge 1020
taataaagat acttgaaata gcttagttta aatataaata gcataataga ttttaggaat 1080
tagtattttg agtttaatta cttattgact tgtaacagtt tttataattc caaggcccat 1140
gaaaaattta atgctttatt agttttaaac ttactatata aatttttcat atgtaaaatt 1200
taatcggtat agttcgatat tttttcaatt tatttttata aaataaaaaa cttaccctaa 1260
ttatcggtac agttatagat ttatataaaa atctacggtt cttcagaaga aacctaaaaa 1320
tcggttcggt gcggacggtt cgatcggttt agtcgatttt caaatattca ttgacactcc 1380
tagttgttgt tataggtaaa aagcagttac agagaggtaa aatataactt aaaaaatcag 1440
ttctaaggaa aaattgactt ttatagtaaa tgactgttat ataaggatgt tgttacagag 1500
```

aggtatgagt gtagttggta aattatgttc ttgacggtgt atgtcacata ttatttatta 1560 aaactagaaa aaacagcgtc aaaactagca aaaatccaac ggacaaaaaa atcggctgaa 1620 tttgatttgg ttccaacatt taaaaaagtt tcagtgagaa agaatcggtg actgttgatg 1680 atataaacaa agggcacatt ggtcaataac cataaaaaat tatatgacag ctacagttgg 1740 tagcatgtgc teagetattg aacaaatcta aagaaggtac atctgtaacc ggaacaccac 1800 ttaaatgact aaattaccct catcagaaag cagatggagt gctacaaata acacactatt 1860 caacaaccat aaataaaacg tgttcagcta ctaaaacaaa tataaataaa tctatgtttg 1920 taagcactcc agccatgtta atggagtgct attgcctgtt aactctcact tataaaatag 1980 tagtagaaaa aatatgaacc aaaacacaac aacatctcaa aatatttgaa qtaacacaga 2040 attttacata caccaaactt ataaatcaag tattttcatt gtaacaaatt ccatgaaaca 2100 tgaaaacaaa gctataatga aattaccaac tcaagcaata aggttggaaa agagccatct 2160 gagatattee ageaatttae atetttttgt ttgattaeae agtgaaggat ettttgtttg 2220 acaactagta aaatgattet tatttgeace ttteaqetat teaqetqett ttacteeaac 2280 cctatagcag aagtaatggc gctcatgctc gttttgtacg ccttccaact tcaagggcga 2340 gctctgtatc gatcttcagg gaaatgtcaa gtgctttcac ggaatgcgtc agggcaccac 2400 tagaaatgta ggtaacacca gtttgtccaa tcttgtgtac tgtttcaagg gtaacatttc 2460 ctgaageete egtateaaae eteceattga teaattetae ageeteetta ageatggata 2520 catcaatatc teegttagat aatggaacaa eeatattgte eageattate etagteaaeg 2580 aagtetttgt ttgagatgea tagtetagaa eeteaegtae ttetteaatt gteetggttt 2640 caacctcaac ccctatttga agtttatttt gctccaaata ctgatccaca gattttagag 2700 ctttgccgac acctccagca gcagatatgt gattgtcttt tatcattacc atatcaaata 2760 agoccattet gtgattette eccecacega teaatacege ecatttatee accaaaegta 2820 atccaggagc agttttccta gtctccaaga tgtaagcagg gtgtgcagca tctgccattt 2880 cettagttag tgtagetatt ceaetcatte tttgeataaa attgagaaca acceteteag 2940 ctataacaat gttgtaagcg tttccttgta ctttgccaaa tttcaagcct ttatgaactt 3000 tategecate atttacatae eactecacet ttaatgaagg atcaaettee gegaatatea 3060 teteageaag tgeaatteet getatgatee egtetteett tgetagaaaa tgageategg 3120 attecatate aagaggaatt gtegeettae aagteacate teetaaatte eeageatett 3180 cagagagtgc aagtttcata actteettta aateataagt tgggtgtget ggtggtttca 3240 cctctaatga ctccactctt qtattcttqq tqqctattqc tqacattttc accaccaacc 3300 ttggagctgt aattgcataa ggatgcactg tagcagtgaa aggaatagct ctaaacatgg 3360 tttttttttg ggggggttgt gaaatgaatt ttgtggaaaa tagtttttgg ggcacatcaa 3420 teetgeggtg acatteggaa tgtttetaae aagaaagata tegttggtee gageettget 3480 ctacatcata geteagtgea taggggeeet gtgegggtge geettagtea agacattgea 3540 gcgagatcat tacaaccact atggcggtgg cgctaaccag ctcgttgatg gttatagccg 3600 aggeactgge ettgetgttg agattatggg cacetttatt ettetgtata etgtettete 3660 egecactgat eccaaaegea atgetagaga tteceatgtt eetgtettgg etecaeteee 3720 cattggettt getgtettea ttgtteacet egecaecatt eeegteaceg geactggeat 3780 accatgttta gagetattee ttteactget acagtgeate ettatgeaat tacageteea 3900 aggttggtgg tgaaaatgtc agcaatagcc accaagaata caagagtgga gtcattagag 3960 gtgaaaccac cagcacaccc aacttatgat ttaaaggaag ttatgaaact tgcactctct 4020 gaagatgctg ggaatttagg agatgtgact tgtaaggcga caattcctct tgatatggaa 4080 teegatgete attttetage aaaggaagae gggateatag eaggaattge aettgetgag 4140 atgatatteg eggaagttga teetteatta aaggtggagt ggtatgtaaa tgatggegat 4200 aaagttcata aaggettgaa atttggeaaa gtacaaggaa aegettacaa eattgttata 4260 gctgagaggg ttgttctcaa ttttatgcaa agaatgagtg gaatagctac actaactaag 4320 gaaatggcag atgctgcaca ccctgcttac atcttggaga ctaggaaaac tgctcctgga 4380 ttacgtttgg tggataaatg ggcggtattg atcggtgggg ggaagaatca cagaatgggc 4440 ttatttgata tggtaatgat aaaagacaat cacatatctg ctgctggagg tgtcggcaaa 4500 gctctaaaat ctgtggatca gtatttggag caaaataaac ttcaaatagg ggttgaggtt 4560 gaaaccagga caattgaaga agtacgtgag gttctagact atgcatctca aacaaagact 4620 tcgttgacta ggataatgct ggacaatatg gttgttccat tatctaacgg agatattgat 4680 gtatccatgc ttaaggaggc tgtagaattg atcaatggga ggtttgatac ggaggcttca 4740 ggaaatgtta cccttgaaac agtacacaag attggacaaa ctggtgttac ctacatttct 4800 agtggtgccc tgacgcattc cgtgaaagca cttgacattt ccctgaagat cgatacagag 4860 ctcgcccttg aagttggaag gcgtacaaaa cgagcatgag cgccattact tctgctatag 4920

```
ggttggagta aaagcagctg aatagctgaa aggtgcaaat aagaatcatt ttactagttg 4980
tcaaacaaaa gateetteae tgtgtaatea aacaaaaaga tgtaaattge tggaatatet 5040
cagatggctc ttttccaacc ttattgcttg agttggtaat ttcattatag ctttgttttc 5100
atgtttcatg gaatttgtta caatgaaaat acttgattta taagtttggt gtatgtaaaa 5160
ttctgtgtta cttcaaatat tttgagatgt tgagctcgtg aaatggcctc tttagttttt 5220
gattgaatca taggggtatt agttttctat ggccgggagt ggtcttcttg cttaattgta 5280
atgqaataac caqaqaqaa ctactgtgtt atctttqaqq aatgttgggc ttttttcgtt 5340
tgaattatca tgaatgaaat tttacttttt cccaatacaa gtttgttttc gtttcttggt 5400
ttttgttatc ccttggttta tgtcttggtt tggcttaaat gattgaagat tacactacct 5460
atgtttctgc tattcctgtt gaagatcaca tttgataata atgcatcgaa tgcattaaag 5520
tttcttattg gctctgtcaa aagtattgaa ggtggatttt tctaattggc aagagaaagt 5580
attaaaqaqq tqatttatta qtacttatat ttttctcaqc atctctcttt caqtqttqqa 5640
getteataaa attaqeaett eaqaqtttea qteqqqaqet qaatteqa
<210> 2
<211> 3600
<212> DNA
<213> Artificial Sequence
<220>
<223> Selection cassette for full-length QTPase RNAi
     construct
<400> 2
tctagaatgt tcqtqcqtca aatgqataaa caaaaaaata gcataagtta gttttgttac 60
tcgagagtta tgtattataa ggtataggga aatgactcaa acataccact gaacttaacg 120
aaacgacgca tatatatact acttaactta acgaaaaagg ggtgagagtg gatgggtgct 180
ggtaaataat gaagggttta tataacgtca cgtgtcaaaa ttcgatagta gtagtttcgt 240
tagttgtaat agcatatatg gcccaaagtt ataatataga taatatgttt atgtccaact 300
attaacgagt gacatagaca gttcattttg tgaagttcaa tgacatattt gagccctttc 360
ccttttatta tctcctttta tttgttctaa taaaagaatg gcatttatta tgtacataga 420
caaataacta ttttctttgg aatataattt gtttatatat tttaaaaatca tgtctcaatt 480
tagtttgttt tgtgcatatt tcaactattc aattttgtcc atatatttat taccttcccc 540
catttacaag cattgaaccg ctttgctcac caaaacttat gcacattgca aaaatatcat 600
gtaaaggttt tatgtatgct gtaattaagg tetgaactca tegtgatttt atttttagge 660
ttcattgacc actaccaaac tctttgatgc tacattttct aattatattg gagttcgatt 720
atateegaat tegegttgeg tagggeeeat tegagggaaa acaeteeeta teaaggattt 780
tttcataccc agagetegaa etcaagacat etggttaagg gaagaacagt etcatecact 840
gcaccatatc cttttgtggt caacaagtaa attttatgta gaaccaaaaa ctatactcga 900
attgataaaa taaatggtgt aaaatattgt tttctttctt acattttgga cagtaaatat 960
gtaggacaat aataattagc gtggggtctt aagaaaatta gcatagattt ccagaaattc 1020
caaatcaacc ggcagttcca ggtttgaaaa ctacaactca ttccgacggt tcaaacttca 1080
aaccatgett getgactegg ettettett ettttteace aagacagage agtagteacg 1140
tgacacccct cacgtgcctc ccccctttat atttcagact gcaaccctac actttcgcta 1200
cattcactac catattettt teactaaqea attttetete etacttttet ttaaacceet 1260
tttttctccc ctaaqccatq qcatctaqat catgttacgt cctgtagaaa ccccaacccg 1320
tgaaatcaaa aaactcgacq qcctgtgggc attcagtctg gatcgcgaaa actgtggaat 1380
tgatcagcgt tggtgggaaa gcgcgttaca agaaagccgg gcaattgctg tgccaggcag 1440
ttttaacgat cagttcgccg atgcagatat tcgtaattat gcgggcaacg tctggtatca 1500
gegegaagte tttatacega aaggttggge aggeeagegt ategtgetge gtttegatge 1560
ggtcactcat tacggcaaag tgtgggtcaa taatcaggaa gtgatggagc atcagggcgg 1620
ctatacgcca tttgaagccq atgtcacgcc gtatgttatt gccgggaaaa gtgtacgtat 1680
caccgtttgt gtgaacaacg aactgaactg gcagactatc ccgccgggaa tggtgattac 1740
cgacgaaaac ggcaaqaaaa aqcaqtctta cttccatgat ttctttaact atgccggaat 1800
ccatcgcagc gtaatgctct acaccacgcc gaacacctgg gtggacgata tcaccgtggt 1860
gacgcatgte gegcaagact gtaaccacge gtetgttgac tggcaggtgg tggccaatgg 1920
tgatgtcagc gttgaactgc gtgatgcgga tcaacaggtg gttgcaactg gacaaggcac 1980
```

tagegggaet ttgeaagtgg tgaateegea eetetggeaa eegggtgaag gttateteta 2040 tgaaetgtge gteaeageea aaageeagae agagtgtgat atetaeeege ttegegtegg 2100

```
catccggtca gtggcagtga agggcgaaca gttcctgatt aaccacaaac cgttctactt 2160
tactggcttt ggtcgtcatg aagatgcgga cttgcgtggc aaaggattcg ataacgtgct 2220
gatggtgcac gaccacgcat taatggactg gattggggcc aactectacc gtacctcgca 2280
ttaccettac getgaagaga tgetegaetg ggeagatgaa catggeateg tggtgattga 2340
tgaaactgct gctgtcggct ttaacctctc tttaggcatt ggtttcgaag cgggcaacaa 2400
gccgaaagaa ctgtacagcg aagaggcagt caacggggaa actcagcaag cgcacttaca 2460
ggcgattaaa gagctgatag cgcgtgacaa aaaccaccca agcgtggtga tgtggagtat 2520
tgccaacgaa ccggataccc gtccgcaagg tgcacgggaa tatttcgcgc cactggcgga 2580
agcaacgcgt aaactcgacc cgacgcgtcc gatcacctgc gtcaatgtaa tgttctgcga 2640
cgctcacacc gataccatca gcgatctctt tgatgtgctg tgcctgaacc gttattacgg 2700
atggtatgtc caaagcggcg atttggaaac ggcagagaag gtactggaaa aagaacttct 2760
ggcctggcag gagaaactgc atcagccgat tatcatcacc gaatacggcg tggatacgtt 2820
ageegggetg cacteaatgt acacegacat gtggagtgaa gagtateagt gtgcatgget 2880
ggatatgtat caccgcgtct ttgatcgcgt cagcgccgtc gtcggtgaac aggtatggaa 2940
tttcgccgat tttgcgacct cgcaaggcat attgcgcgtt ggcggtaaca agaaagggat 3000
cttcactcgc gaccgcaaac cgaagtcggc ggcttttctg ctgcaaaaac gctggactgg 3060
catgaacttc ggtgaaaaac cgcagcaggg aggcaaacaa tgagagctcg tgaaatggcc 3120
tetttagttt ttgattgaat cataggggta ttagttttet atggeeggga gtggtettet 3180
tgcttaattg taatggaata accagagagg aactactgtg ttatctttga ggaatgttgg 3240
getttttteg tttgaattat catgaatgaa attttaettt tteecaatae aagtttgttt 3300
tegtttettg gtttttgtta tecettggtt tatgtettgg tttggettaa atgattgaag 3360
attacactac ctatgtttct gctattcctg ttgaagatca catttgataa taatgcatcg 3420
aatgcattaa agtttcttat tggctctgtc aaaagtattg aaggtggatt tttctaattg 3480
gcaagagaaa gtattaaaga ggtgatttat tagtacttat atttttctca gcatctctct 3540
ttcagtgttg gagcttcata aaattagcac ttcagagttt cagtcgggag ctgaattcga 3600
<210> 3
<211> 4134
<212> DNA
<213> Artificial Sequence
<220>
<223> Partial QTPase RNAi construct
ctcgaggatc taaattgtga gttcaatctc ttccctattg gattgattat cctttcttt 60
cttccaattt gtgtttcttt ttgcctaatt tattgtgtta tcccctttat cctattttgt 120
ttetttaett atttatttge ttetatgtet ttgtacaaag atttaaaete tatggeacat 180
attttaaagt tgttagaaaa taaattettt caagattgat gaaagaaett tttaattgta 240
gatatttcgt agattttatt ctcttactac caatataacg cttgaattga cgaaaatttg 300
tgtccaaata tctagcaaaa aggtatccaa tgaaaatata tcatatgtga tcttcaaatc 360
ttgtgtctta tgcaaqattq atactttqtt caatggaaga gattqtqtgc atatttttaa 420
aatttttatt agtaataaag attctatata gctgttatag agggataatt ttacaaagaa 480
cactataaat atgattgttg ttgttagggt gtcaatggtt cggttcgact ggttatttta 540
taaaatttgt accataccat ttttttcgat attctatttt gtataaccaa aattagactt 600
ttcgaaatcg tcccaatcat gtcggtttca cttcggtatc ggtaccgttc ggttaatttt 660
cattttttt taaatgtcat taaaattcac tagtaaaaat agaatgcaat aacatacgtt 720
cttttatagg acttagcaaa agctctctag acatttttac tgtttaaagg ataatgaatt 780
aaaaaacatg aaagatggct agagtataga tacacaacta ttcgacagca acgtaaaaga 840
aaccaagtaa aagcaaagaa aatataaatc acacgagtgg aaagatatta accaagttgg 900
gattcaagaa taaagtctat attaaatatt caaaaagata aatttaaata atatgaaagg 960
aaacatattc aatacattgt agtttgctac tcataatcgc tagaatactt tgtgccttgc 1020
taataaagat acttgaaata gcttagttta aatataaata gcataataga ttttaggaat 1080
tagtattttg agtttaatta cttattgact tgtaacagtt tttataattc caaggcccat 1140
```

```
qaaaaattta atqctttatt aqttttaaac ttactatata aatttttcat atqtaaaatt 1200
taatcqqtat aqttcqatat tttttcaatt tattttata aaataaaaaa cttaccctaa 1260
ttatcggtac agttatagat ttatataaaa atctacggtt cttcagaaga aacctaaaaa 1320
teggtteggt geggaeggtt egateggttt agtegatttt caaatattea ttgacaetee 1380
tagttgttgt tataggtaaa aagcagttac agagaggtaa aatataactt aaaaaatcag 1440
ttctaaggaa aaattgactt ttatagtaaa tgactgttat ataaggatgt tgttacagag 1500
aggtatgagt gtagttggta aattatgttc ttgacggtgt atgtcacata ttatttatta 1560
aaactaqaaa aaacagcqtc aaaactaqca aaaatccaac ggacaaaaaa atcqqctqaa 1620
tttgatttgg ttccaacatt taaaaaagtt tcagtgagaa agaatcggtg actgttgatg 1680
atataaacaa agggcacatt ggtcaataac cataaaaaat tatatgacag ctacagttgg 1740
tagcatgtgc tcagctattg aacaaatcta aagaaggtac atctgtaacc ggaacaccac 1800
ttaaatgact aaattaccct catcagaaag cagatggagt gctacaaata acacactatt 1860
caacaaccat aaataaacq tqttcaqcta ctaaaacaaa tataaataaa tctatqtttq 1920
taagcactcc agccatgtta atggagtgct attgcctgtt aactctcact tataaaatag 1980
tagtagaaaa aatatgaacc aaaacacaac tttatcgcca tcatttacat accactccac 2040
ctttaatgaa ggatcaactt ccgcgaatat catctcagca agtgcaattc ctgctatgat 2100
cccgtcttcc tttgctagaa aatgagcatc ggattccata tcaagaggaa ttgtcgcctt 2160
acaagtcaca teteetaaat teecageate tteagagagt geaagtttea taaetteett 2220
taaatcataa gttgggtgtg ctggtggttt cacctctaat gactccactc ttgtattctt 2280
ggtggctatt gctgacattt tcaccaccaa ccttggaget gtaattgcat aaggatgcac 2340
tqtaqcaqtq aaaqqaataq ctctaaacat qtccqtcqct tctcttccat ttcttctcat 2400
tttcgatttt gattcttatt tctttccagt agctcctgct ctgtgaattt ctccgctcac 2460
gatagatetg ettataetee ttaeatteaa eettagatet ggtetegatt etetgtttet 2520
ctqttttttt cttttqqtcq aqaatctqat qtttqtttat qttctqtcac cattaataat 2580
aatqaactct ctcattcata caatqattaq tttctctcqt ctacaaaacq atatqttqca 2640
ttttcacttt tcttctttt ttctaagatg atttgctttg accaatttgt ttagatcttt 2700
attttatttt attttctggt gggttggtgg aaattgaaaa aaaaaaaaac agcataaatt 2760
gttatttgtt aatgtattca ttttttggct atttgttctg ggtaaaaatc tgcttctact 2820
attgaatett teetggattt tttaeteeta ttgggttttt atagtaaaaa taeataaa 2880
aaggaaaaca aaagttttat agattetett aaaceeetta egataaaagt tggaateaaa 2940
ataattcagg atcagatget etttgattga ttcagatgeg attacagttg catggcaaat 3000
tttctagatc cgtcgtcaca ttttattttc tgtttaaata tctaaatctg atatatgatg 3060
tegacaaatt etgqtqqett atacateact teaactqttt tettttqqet ttqtttqtea 3120
acttggtttt caatacqatt tqtqatttcq atcgctqaat ttttaataca aqcaaactqa 3180
tgttaaccac aagcaagaga tgtgacctgc cttattaaca tcgtattact tactactagt 3240
cgtattctca acgcaatcgt tittgtattt ctcacattat gccgcttctc tactctttat 3300
teettttggt ceaegeattt tetatttgtg geaateeett teaeaacetg attteeeact 3360
ttqqatcatt tqtctqaaqa ctctcttqaa tcqttaccac ttqtttcttq tqcatqctct 3420
gttttttaga attaatgata aaactattcc atagtcttga gttttcagct tgttgattct 3480
tttgcttttg gttttctgca qatgtttaga gctattcctt tcactgctac agtgcatcct 3540
tatgcaatta cagctccaag gttggtggtg aaaatgtcag caatagccac caagaataca 3600
agagtggagt cattagaggt gaaaccacca gcacacccaa cttatgattt aaaggaagtt 3660
atgaaacttg cactetetga agatgetggg aatttaggag atgtgaettg taaggegaca 3720
attectettg atatggaate egatgeteat tttetageaa aggaagaegg gateatagea 3780
ggaattgcac ttgctgagat gatattcgcg gaagttgatc cttcattaaa ggtggagtgg 3840
tatgtaaatg atggcgataa agcaagtgtg ttgcctttgt gtggaaatga agaggtactt 3900
gcgaggactt tgcgtttatc agtttatgtg tttgtatatc tatttgatcc agttattatg 3960
gattatatac gettqaaact cattttaage cattgttatt gaacgtttat caaatacttt 4020
attatgecaa geaagteaaa cacatgettg ttgattgaaa teaagetata gaaatetett 4080
cttcacatac agcagtttag attcacaata caacaagcga aacgataaag tttc
                                                                  4134
```

<210> 4

<211> 3387

<212> DNA

<213> Artificial Sequence

<220>

<223> Selection cassette for partial length QTPase RNAi construct

<400> 4 cgttttgacg agttcggatg tagtagtagc cattatttaa tgtacatact aatcgtgaat 60 agtgaatatg atgaaacatt gtatcttatt gtataaatat ccataaacac atcatgaaag 120 acactttett teaeggtetg aattaattat gatacaatte taatagaaaa egaattaaat 180 tacgttgaat tqtatgaaat ctaattqaac aaqccaacca cqacqacqac taacqttqcc 240 tggattgact cggtttaagt taaccactaa aaaaacggag ctgtcatgta acacgcggat 300 cgagcaggtc acagtcatga agccatcaaa gcaaaagaac taatccaagg gctgagatga 360 ttaattagtt taaaaattag ttaacacgag ggaaaaggct gtctgacagc caggtcacgt 420 tatetttaee tgtggtegaa atgattegtg tetgtegatt ttaattattt ttttgaaagg 480 CCgaaaataa aqttqtaaqa qataaacccq cctatataaa ttcatatatt ttctctccqc 540 tttgaattgt ctcgttgtcc tcctcacttt catcgqccqt ttttgaatct ccqqcqactt 600 gacagagaag aacaaggaag aagactaaga gagaaagtaa gagataatcc aggagattca 660 ttctccgttt tgaatcttcc tcaatctcat cttcttccgc tctttctttc caaggtaata 720 ggaactttct ggatctactt tatttgctgg atctcgatct tgttttctca atttccttga 780 gatctggaat tcgtttaatt tggatctgtg aacctccact aaatcttttg gttttactag 840 aatcgatcta agttgaccga tcagttagct cgattatagc taccagaatt tggcttgacc 900 ttgatggaga qatccatqtt catqttacct qqqaaatqat ttqtatatqt qaattqaaat 960 ctgaactgtt gaagttagat tgaatctgaa cactgtcaat gttagattga atctgaacac 1020 tgtttaagtt agatgaagtt tgtgtataga ttcttcgaaa ctttaggatt tgtagtgtcg 1080 tacgttgaac agaaagctat ttctgattca atcagggttt atttgactgt attgaactct 1140 ttttgtgtgt ttgcagetca tatggttgtg tttgggaatg tttctqcqc qaatttqcct 1200 tatcaaaacg ggtttttgga ggcactttca tctggaggtt gtgaactaat gggacatagc 1260 tttagggttc ccacttetca agegettaag acaagaacaa qqaqqaqqaq tactqetqqt 1320 cetttgcagg tagtttgtgt ggatattcca aggccagage tagagaacac tgtcaatttc 1380 ttggaagctg ctagtttatc tgcatccttc cgtagtgctc ctcgtcctgc taagcctttg 1440 aaagttgtaa ttgctggtgc tggattggct ggattgtcaa ctgcaaagta cctggctgat 1500 gcaggccaca aacctctgtt gcttgaagca agagatgttc ttggtggaaa gatagctgca 1560 tggaaggatg aagatgggga ctggtatgag actggtttac atattttctt cggtgcttat 1620 ccgaatgtgc agaatttatt tggagaactt gggatcaatg atcggttgca gtggaaggaa 1680 cactecatga tttttgctat gccaagtaaa cctggagaat ttaqtaqatt tqacttccca 1740 gatgtcctac cagcaccctt aaatggtatt tgggctattt tgcggaacaa cgagatgctg 1800 acatggccag agaaaataaa gtttgctatt ggacttttgc cagccatggt cggcggtcag 1860 gcttatgttg aggcccaaga tggtttatca gtcaaagaat ggatggaaaa gcagggagta 1920 cctgagcgcg tgaccgacga ggtgtttatt gccatgtcaa aggcgctaaa ctttataaac 1980 cctgatgaac tgtcaatgca atgcattttg atagctttga accggtttct tcaggaaaaa 2040 catggttcca agatggcatt cttggatggt aatectccgg aaaggctttg tatgccagta 2100 gtggatcata ttcqatcact aqqtqqqqaa qtqcaactta attctaqqat aaaqaaaatt 2160 gageteaatg acgatggeac ggttaagagt ttettaetea etaatggaag caetgtegaa 2220 ggagacgett atgtgtttge egeteeagte gatateetga ageteetttt accagateee 2280 tggaaagaaa taccgtactt caagaaattg gataaattag ttggagtacc agttattaat 2340 gttcatatat ggtttgatcg aaaactgaag aacacatatg atcacctact ctttagcaga 2400 agtaaccttc tgagcgtgta tgccgacatg tccttaactt gtaaggaata ttacgatcct 2460 aaccggtcaa tgctggagct agtatttgca ccagcagagg aatggatatc acggactgat 2520 tetgacatea tagatgeaac aatgaaagaa ettgagaaac tetteeetga tgaaatetea 2580 gctgaccaaa gcaaaqctaa aattctgaag taccatqtcq ttaaqactcc aaqatctqqq 2640 tacaagacca teceaaactg tgaaecatgt egteetetae aaagateace tattgaagga 2700 ttctacttag ctggagatta cacaaaacag aagtacttag cttccatgga aggcgctgtc 2760 ctctctggca aattctgctc tcagtctatt gttcaggatt acgagctact ggctgcgtct 2820 ggaccaagaa agttgtegga ggcaacagta teateateat gagaaaaggg egaattegtt 2880 aaccgcagac gagctcgtga aatggcctct ttagtttttg attgaatcat aggggtatta 2940 gttttctatg gccgggagtg gtcttcttgc ttaattgtaa tggaataacc agagaggaac 3000 tactgtgtta tetttqaqqa atqttqqqet tttttcqttt qaattateat qaatqaaatt 3060 ttactttttc ccaatacaag tttgttttcg tttcttggtt tttgttatcc cttggtttat 3120 gtcttggttt ggcttaaatg attgaagatt acactaccta tgtttctgct attcctgttg 3180

```
aagatcacat ttgataataa tgcatcgaat gcattaaagt ttcttattgg ctctgtcaaa 3240
agtattgaag gtggattttt ctaattggca agagaaagta ttaaagaggt gatttattag 3300
tacttatatt tttctcagca tctcttttc agtgttggag cttcataaaa ttagcacttc 3360
                                                                  3387
agagtttcag tcgggagctg aattcga
<210> 5
<211> 3458
<212> DNA
<213> Arabidopsis thaliana
<400> 5
atggttgtgt ttgggaatgt ttctgcggcg aatttgcctt atcaaaacgg gtttttggag 60
gcactttcat ctggaggttg tgaactaatg ggacatagct ttagggttcc cacttctcaa 120
gcgcttaaga caagaacaaq qaqqaggaqt actgctggtc ctttgcaggt agtttgtgtg 180
qatattecaa qqccaqaqct aqaqaacact qtcaatttet tqqaaqetqc taqtttatet 240
geatectice qtaqtqctcc teqtectqct aagectitga aagttqtaat tqctqqtqct 300
ggattggctg gattgtcaac tgcaaagtac ctggctgatg caggccacaa acctctgttg 360
cttgaagcaa gagatgttct tggtggaaag atagctgcat ggaaggatga agatggggac 420
tggtatgaga ctggtttaca tattttcttc ggtgcttatc cgaatgtgca gaatttattt 480
ggagaacttg ggatcaatga tcggttgcag tggaaggaac actccatgat ttttgctatg 540
ccaagtaaac ctggagaatt tagtagattt gacttcccag atgtcctacc agcaccctta 600
aatggtattt qqqctatttt qcqqaacaac gagatgctga catgqccaga gaaaataaag 660
tttgetattg gaettttgee ageeatggte ggeggteagg ettatgttga ggeecaagat 720
ggtttatcag tcaaagaatg gatggaaaag cagggagtac ctgagcgcgt gaccgacgag 780
gtgtttattg ccatgtcaaa ggcgctaaac tttataaacc ctgatgaact gtcaatgcaa 840
tgcattttga taqctttgaa ccqqtttctt caggaaaaac atgqttccaa gatggcattc 900
ttggatggta atcctccgga aaggctttgt atgccagtag tggatcatat tcgatcacta 960
ggtggggaag tgcaacttaa ttctaggata aagaaaattg agctcaatga cgatggcacg 1020
gttaagagtt tettaeteae taatggaage aetgtegaag gagaegetta tgtgtttgee 1080
gctccagtcg atatcctgaa gctcctttta ccagatccct ggaaagaaat accgtacttc 1140
aagaaattgg ataaattagt tggagtacca gttattaatg ttcatatatg gtttgatcga 1200
aaactgaaga acacatatga teacetacte tttagcagaa gtaacettet gagegtgtat 1260
gccgacatgt cettaaettg taaggaatat tacgateeta accggteaat getggageta 1320
gtatttgcac caqcaqatqq ttqtgtttqg gaatgtttct gcggcgaatt tgccttatca 1380
aaacgggttt ttggaggcac tttcatctgg aggttgtgaa ctaatgggac atagctttag 1440
ggttcccact tctcaagcgc ttaagacaag aacaaggagg aggagtactg ctggtccttt 1500
gcaggtagtt tgtgtggata ttccaaggcc agagctagag aacactgtca atttcttgga 1560
agetgetagt ttatetgeat cetteegtag tgeteetegt cetgetaage etttgaaagt 1620
tgtaattgct ggtgctggat tggctggatt gtcaactgca aagtacctgg ctgatgcagg 1680
ccacaaacct ctqttqcttq aaqcaaqaqa tqttcttqqt qqaaaqataq ctqcatqgaa 1740
ggatgaagat ggggactggt atgagactgg tttacatatt ttcttcggtg cttatccgaa 1800
tgtqcaqaat ttatttqqaq aacttqqqat caatqatcqq ttgcaqtqqa aqqaacactc 1860
catgattttt gctatgccaa gtaaacctgg agaatttagt agatttgact tcccagatgt 1920
cctaccagca cccttaaatg gtatttgggc tattttgcgg aacaacgaga tgctgacatg 1980
gccagagaaa ataaagtttg ctattggact tttgccagcc atggtcggcg gtcaggctta 2040
tgttgaggcc caagatggtt tatcagtcaa agaatggatg gaaaagcagg gagtacctga 2100
gegegtgace gaegaggtgt ttattgeeat gteaaaggeg etaaaettta taaaceetga 2160
tgaactgtca atgcaatgca ttttgatagc tttgaaccgg tttcttcagg aaaaacatgg 2220
ttccaagatq qcattcttqq atqqtaatcc tccgqaaagg ctttqtatqc cagtagtgga 2280
tcatattcga tcactaggtg gggaagtgca acttaattct aggataaaga aaattgagct 2340
caatgacgat ggcacggtta agagtttctt actcactaat ggaagcactg tcgaaggaga 2400
egettatgtg tttgeegete eagtegatat eetgaagete ettttaeeag ateeetggaa 2460
agaaataccg tacttcaaga aattggataa attagttgga gtaccagtta ttaatgttca 2520
tatatggttt gatcgaaaac tgaagaacac atatgatcac ctactcttta gcagaagtaa 2580
ccttctqaqc qtqtatqccq acatqtcctt aacttqtaaq qaatattacq atcctaaccq 2640
gtcaatgctg qaqctaqtat ttgcaccagc agaggaatgg atatcacgga ctgattctga 2700
catcatagat gcaacaatga aagaacttga gaaactcttc cctgatgaaa tctcagctga 2760
```

ccaaagcaaa	gctaaaattc	tgaagtacca	tgtcgttaag	actccaagat	ctgggtacaa	2820
gaccatccca	aactgtgaac	catgtcgtcc	tctacaaaga	tcacctattq	aaggattcta	2880
cttagctgga	gattacacaa	aacagaagta	cttagcttcc	atggaaggcg	ctgtcctctc	2940
tggcaaattc	tgctctcagt	ctattgttca	ggattacgag	ctactggctg	cgtctggacc	3000
aagaaagttg	tcggaggcaa	cagtatcatc	atcatgagaa	aagggcgaat	tcgttaaccg	3060
cagacaggaa	tggatatcac	ggactgattc	tgacatcata	gatgcaacaa	tgaaagaact	3120
tgagaaactc	ttccctgatg	aaatctcagc	tgaccaaagc	aaagctaaaa	ttctgaagta	3180
ccatgtcgtt	aagactccaa	gatctgggta	caagaccatc	ccaaactgtg	aaccatgtcg	3240
teetetacaa	agatcaccta	ttgaaggatt	ctacttagct	ggagattaca	caaaacagaa	3300
gtacttagct	tccatggaag	gcgctgtcct	ctctggcaaa	ttctgctctc	agtctattgt	3360
tcaggattac	gagctactgg	ctgcgtctgg	accaagaaag	ttgtcggagg	caacagtatc	3420
atcatcatga	gaaaagggcg	aattcgttaa	ccgcagac		_	3458